

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION  
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété  
Intellectuelle  
Bureau international



(43) Date de la publication internationale  
4 novembre 2004 (04.11.2004)

PCT

(10) Numéro de publication internationale  
**WO 2004/095840 A2**

- (51) Classification internationale des brevets<sup>7</sup> : **H04N 7/167**
- (21) Numéro de la demande internationale : **PCT/FR2004/000947**
- (22) Date de dépôt international : 16 avril 2004 (16.04.2004)
- (25) Langue de dépôt : français
- (26) Langue de publication : français
- (30) Données relatives à la priorité :  
03/50110 16 avril 2003 (16.04.2003) FR
- (71) Déposant (pour tous les États désignés sauf US) : **MEDI-ALIVE [FR/FR]**; 111, avenue Victor Hugo, F-75116 Paris (FR).
- (72) Inventeurs; et
- (73) Inventeurs/Déposants (pour US seulement) : **LECOMTE, Daniel [FR/FR]**; 157, rue de La Pompe, F-75116 Paris (FR). **CAPOROSSI, Jérôme [FR/FR]**; 7, rue du 8 mai 1945, F-92340 Bourg-La-Reine (FR).
- (74) Mandataires : **BREESE, Pierre etc.**; Breesé-Majerowicz, 3, avenue de l'Opéra, F-75001 Paris (FR).
- (81) États désignés (sauf indication contraire, pour tout titre de protection nationale disponible) : AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH,

[Suite sur la page suivante]

(54) Title: SCRAMBLING, DESCRAMBLING AND SECURE DISTRIBUTION OF AUDIO-VISUAL SEQUENCES FROM VIDEO ENCODERS BASED ON WAVELET PROCESSING

(54) Titre : EMBROUILLAGE ET DESEMBROUILLAGE ET DISTRIBUTION SECURISEE DE SEQUENCES AUDIOVISUELLES ISSUES DE CODEURS VIDEOS BASES SUR UN TRAITEMENT PAR ONDELETTES

(57) Abstract: The invention relates to a method and system which are used for the visual scrambling of a video sequence and for the reconstruction (descrambling) of the original content thereof from a digital video stream obtained from encoding based on a wavelet transform. The method for the secure distribution of video sequences according to a digital stream format stems from wavelet-based coding, said format consisting of frames comprising blocks containing wavelet coefficients which describe visual elements. According to the invention, before transmission to the client equipment, the stream is analysed in order to generate (i) a modified main stream through the removal and replacement of certain pieces of information encoding the original stream and having the same format as said original stream, and (ii) a complementary piece of information with any format, comprising digital information which encode the original stream and which can be used to reconstruct the aforementioned modified frames. The modified main stream and the complementary information thus generated are then transmitted separately from the server to the destination equipment and, subsequently, a synthesis of a stream with the nominal format is calculated on the destination equipment as a function of the modified main stream and the complementary information. The invention also relates to a system comprising at least one multimedia server containing the original video sequences, a video stream analysis device, a device for separating the original video stream into a modified main stream and a complementary piece of information according to the aforementioned analysis and at least one device on the destination equipment for the reconstruction of the video stream as a function of the modified main stream and the complementary information.

(57) Abrégé : La présente invention propose un procédé et un système permettant d'embrouiller visuellement une séquence vidéo et de recomposer (désembrouiller) son contenu original à partir d'un flux vidéo numérique obtenu par un encodage reposant sur une transformée en ondelettes. Le procédé pour la distribution sécurisée de séquences vidéos selon un format de flux numérique est issu d'un encodage basé ondelettes, constitué de trames (frames) comprenant des blocs contenant des coefficients d'ondelettes décrivant les éléments visuels. On procède, avant la transmission à l'équipement client, à une analyse du flux pour générer un flux principal modifié, par suppression et remplacement de certaines informations codant le flux original et présentant le format du flux original, et une information complémentaire d'un format quelconque, comportant les informations numériques codant le flux original, aptes à permettre la reconstruction desdites trames modifiées, puis on transmet séparément ledit flux principal modifié et ladite information complémentaire ainsi générés depuis le serveur vers l'équipement destinataire, et on calcule sur l'équipement destinataire une synthèse d'un flux au format nominal en fonction dudit flux principal modifié et de ladite information complémentaire. L'invention concerne également un système comportant au moins un serveur multimédia contenant les séquences vidéos originales, un dispositif d'analyse du flux vidéo, un dispositif de séparation du flux vidéo original en un flux principal modifié et en une information complémentaire en fonction de ladite analyse et au moins un dispositif sur l'équipement destinataire pour la reconstruction du flux vidéo en fonction dudit flux principal modifié et de ladite information complémentaire.

WO 2004/095840 A2



PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN,  
TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) États désignés (sauf indication contraire, pour tout titre de protection régionale disponible) : ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Publiée :

— sans rapport de recherche internationale, sera republiée dès réception de ce rapport

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

EMBROUILLAGE, DESEMBROUILLAGE ET DISTRIBUTION  
SECURISEE DE SEQUENCES AUDIOVISUELLES ISSUES DE CODEURS  
VIDEOS BASES SUR UN TRAITEMENT PAR ONDELETTES

La présente invention se rapporte au domaine du  
5 traitement de séquences vidéos encodées à l'aide de  
codeurs vidéos basés sur la technologie des ondelettes.

On se propose dans la présente invention de fournir  
un procédé et un système permettant d'embrouiller  
visuellement une séquence vidéo et de recomposer  
10 ultérieurement (désembrouiller) son contenu original à  
partir d'un flux vidéo numérique obtenu par un encodage  
reposant sur une transformée en ondelettes.

La présente invention se rapporte plus  
particulièrement à un dispositif capable de transmettre de  
15 façon sécurisée un ensemble de flux vidéos de haute  
qualité visuelle vers un écran de visualisation type écran  
de télévision et/ou pour être enregistré sur le disque dur  
ou sur tout autre support d'enregistrement d'un boîtier  
reliant le réseau de télétransmission à l'écran de  
20 visualisation tel qu'un écran de télévision ou un moniteur  
d'ordinateur personnel, tout en préservant la qualité  
audiovisuelle mais en évitant toute utilisation  
frauduleuse comme la possibilité de faire des copies  
pirates de films ou de programmes audiovisuels enregistrés  
25 sur le disque dur ou tout autre support d'enregistrement  
du boîtier décodeur. L'invention concerne également un  
système client-serveur, le serveur fournissant le flux  
permettant le visionnage du film vidéo numérique sécurisé  
et le client lisant et affichant le flux audiovisuel  
30 numérique.

Avec les solutions actuelles, il est possible de  
transmettre des films et des programmes audiovisuels sous  
forme numérique via des réseaux de diffusion de type

hertzien, câble, satellite, etc. ou via des réseaux de télécommunication type DSL (Digital Subscriber Line) ou BLR (boucle locale radio) ou via des réseaux DAB (Digital Audio Broadcasting), etc. Par ailleurs, pour éviter le piratage des œuvres ainsi diffusées, ces dernières sont souvent cryptées ou brouillées par divers moyens bien connus de l'homme de l'art.

Concernant le traitement de séquences vidéo encodées avec la technologie des ondelettes, l'art antérieur connaît le brevet référencé US 6370197 et intitulé « Video Compression Scheme using wavelets » dans lequel les auteurs détaillent une méthode de codage d'une séquence vidéo basée sur une transformée en ondelettes et générant un flux numérique emboîté. Cet art antérieur ne propose aucune méthode de protection du flux et/ou d'embrouillage de la séquence vidéo.

On connaît également la demande de brevet EP0734164 présentant un procédé et un dispositif pour augmenter l'efficacité de codage, effectué par des encodeurs vidéo basés sur la quantification vectorielle classifiée en

5 optimisant le codage de manière à ne pas avoir à transmettre l'information de classification dans le flux binaire encodé. Cet art antérieur s'applique à des flux vidéos issus de la transformée en DCT ou en ondelettes. Dans ce but, le signal vidéo entrant est divisé en une

10 pluralité de sous-bandes, par exemple les coefficients DC sont rangés dans une sous-bande et les coefficients AC dans des sous-bandes restantes, suivi d'un formatage en blocks de taille identique, chaque block incluant un coefficient DC et une multitude de coefficients AC. Un

15 signal de sélection est ensuite généré, représentant la classe de quantification vectorielle correspondante à chaque block assemblé. Cette étape est suivie de classification pour la quantification vectorielle par la

génération de paramètres relatifs à l'évolution des coefficients DC dans le sens horizontal et vertical, et de l'encodage entropique différentiel des coefficients DC relatifs aux blocks assemblés pour générer un premier  
5 signal vidéo encodé. Les coefficients AC sont classifiés et encodés séparément à l'aide d'encodage entropique en fonction de l'information de sélection pour générer un second signal vidéo encodé. Les deux signaux ainsi générés sont formatés pour la transmission. Pour le  
10 décodage, aucune information de classification n'est transmise, elle est reconstruite à partir des coefficients DC encodés et transmis au décodeur. Cette solution concerne un procédé relatif à la compression numérique et l'encodage de flux vidéos issus de la transformée en DCT  
15 et ondelettes. La description du procédé indique les étapes à appliquer pour mettre en œuvre une quantification vectorielle classifiée, augmentant la compression et l'efficacité de l'encodage. Un seul flux est transmis vers le récepteur. Le problème technique et l'objectif posés  
20 dans ce document sont d'optimiser le format numérique et il vise l'obtention d'un flux numérique formaté à la sortie d'un encodeur numérique. Le procédé décrit dans cette solution ne permet pas de sécuriser le flux vidéo et n'apporte aucune protection contre des utilisations  
25 illicites de flux vidéos issus de l'encodage en ondelettes.

En ce qui concerne la protection d'images codées en ondelettes, l'art antérieur connaît le document EP 1 033 880 qui est relatif à un procédé et un dispositif de  
30 protection par des modifications appliquées aux coefficients spatio-fréquentiels. Ces modifications sont de type : modification des bits de signe des coefficients, modification des bits d'amélioration des coefficients,

choix des coefficients appropriés appartenant à une sous-bande fréquentielle pour les permuter, rotation d'un block regroupant des coefficients fréquents rangés par ordre croissant, tout en essayant de respecter au maximum les propriétés statistiques et l'entropie du signal original. Chaque type de modifications est conditionné à l'aide d'une clé. Les données ainsi protégées sont ensuite passées par un encodeur entropique et un bitstream conforme à la norme est généré. Cet art antérieur représente une solution de cryptage à l'aide de clés, par conséquent, un seul flux est transmis vers le récepteur et tous les éléments constituant le flux original se trouvent à l'intérieur du flux protégé. Ce document concerne une solution ne répondant pas de manière satisfaisante à la protection du flux vidéo transmis. De plus, suite aux modifications avant l'encodage entropique, les propriétés statistiques sont modifiées et la taille du flux et le débit augmentent. Par conséquent, cet art antérieur ne répond pas aux objectifs de haute sécurité garantissant un procédé sans perte, objet de la présente invention.

Une autre référence de l'art antérieur est le document WO 00 31964 A relatif à une méthode et à un équipement pour le cryptage partiel d'images dans le but de les protéger et d'optimiser la place de stockage. Une première partie de l'image est compressée à faible qualité, sans cryptage, et une seconde partie de l'image est cryptée. Quand la première et la deuxième partie sont réunies, on obtient l'image de qualité maximale. La deuxième partie est cryptée et comprend deux sections, cryptées de manière différente. Le décryptage de la première section et sa combinaison avec la première partie restitue l'image initiale avec une qualité moyenne. Le décryptage de la seconde section, sa combinaison avec la première section et la première partie restituent l'image

originale à qualité maximale. L'image peut également être partitionnée en multiples sections indépendantes, chaque section étant cryptée avec sa propre méthode et sa propre clé. La méthode de protection dans cet art antérieur est

5 le cryptage, par conséquent, tous les éléments originaux du flux restent à l'intérieur du flux protégé, la restitution du contenu entier à partir du seul flux protégé est possible dans le cas où une personne mal intentionnée trouve ou simule les clés de cryptage. Comme

10 pour le document précédent, cette solution n'apporte pas une sécurité satisfaisante contre le piratage du flux vidéo. Aussi, la taille du flux protégé est différente de la taille du flux original. Cet art antérieur ne résout donc pas le problème de haute sécurité, tout en procurant

15 une fine granularité dans la qualité des séquences vidéos reconstituées, traité dans la présente invention.

Parmi l'art antérieur concernant la distribution sécurisée de flux audiovisuels organisés en multicouches basés sur le principe client - serveur, le brevet US 2001/0053222 A1 propose un procédé et un système pour la protection de flux vidéos encodés selon la norme MPEG-4. Le flux audiovisuel est composé de plusieurs objets audio et vidéos, gérés par une composition scénique. Un des objets du flux vidéo est crypté à l'aide d'une clé générée en quatre étapes de cryptage qui est renouvelée périodiquement. Les objets protégés sont les objets vidéos. L'objet crypté est multiplexé avec les autres objets et le flux entier est envoyé à l'utilisateur. Le flux MPEG-4 est recomposé sur l'équipement destinataire par le module de décryptage, qui reconstitue le flux vidéo original, à partir du flux vidéo crypté, et en régénérant la clé de cryptage à partir d'informations de cryptage envoyées au préalable et à partir d'informations contenues dans le flux crypté. Etant donné que tout le contenu

protégé des objets vidéo se trouve dans le flux envoyé à l'utilisateur, une personne mal intentionnée qui retrouve les clés de cryptage pourra décrypter ce contenu protégé et le visualiser ou diffuser. Cet art antérieur ne résout donc pas entièrement le problème de la sécurisation du flux vidéo.

A l'inverse de la plupart de ces systèmes « classiques » de protection, le procédé conforme à l'invention assure un haut niveau de protection tout en réduisant le volume d'information nécessaire pour avoir accès au contenu original à partir du contenu protégé.

La protection, réalisée de façon conforme à l'invention, est basée sur le principe de suppression et de remplacement de certaines informations codant le signal visuel original par une méthode quelconque, soit : substitution, modification, permutation ou déplacement de l'information. Cette protection est également basée sur la connaissance de la structure du flux binaire à la sortie de l'encodeur visuel basé ondelettes.

La présente invention concerne le principe général d'un procédé de sécurisation d'un flux audiovisuel. L'objectif est d'autoriser les services de vidéo à la demande et à la carte à travers tous les réseaux de diffusion et l'enregistrement local dans le boîtier décodeur numérique de l'utilisateur, ainsi que la visualisation en direct des chaînes de télévision. La solution consiste à extraire et conserver en permanence à l'extérieur de l'habitation de l'utilisateur, en fait dans le réseau de diffusion et de transmission, une partie du flux audiovisuel numérique enregistré chez le client ou diffusé en direct, cette partie étant primordiale pour visualiser ledit flux audiovisuel numérique sur un écran de télévision ou de type moniteur, mais étant d'un volume très faible par rapport au volume total du flux



audiovisuel numérique enregistré chez l'utilisateur ou reçu en temps réel. La partie manquante sera transmise via le réseau de diffusion ou de transmission au moment de la visualisation dudit flux audiovisuel numérique.

Le flux audiovisuel numérique étant modifié et séparé en deux parties, la plus grande partie du flux audiovisuel modifié, appelé «flux principal modifié» sera donc transmise via un réseau de diffusion classique alors que la partie manquante appelée «information complémentaire» sera envoyée à la demande via un réseau de télécommunication bande étroite comme les réseaux téléphoniques classiques ou les réseaux cellulaires de type GSM, GPRS ou UMTS ou en utilisant une petite partie d'un réseau de type DSL ou BLR, ou encore en utilisant un sous-ensemble de la bande passante partagée sur un réseau câblé. Le flux audiovisuel numérique original est reconstitué sur l'équipement destinataire (décodeur) par un module de synthèse à partir du flux principal modifié et de l'information complémentaire.

- 5 L'invention réalise un système de protection, comprenant un module d'analyse - embrouillage et de désembrouillage basé sur un format numérique issu de l'encodage d'un flux vidéo basé sur des transformées en ondelettes. Le module d'analyse et d'embrouillage proposé
- 10 par l'invention repose sur la substitution par des « leurres » ou la modification d'une partie des coefficients issus de la transformation en ondelettes. Le fait d'avoir enlevé et substitué une partie des données originales du flux numérique d'origine lors de la
- 15 génération du flux principal modifié, ne permet pas la restitution dudit flux d'origine à partir des seules données dudit flux principal modifié. De par les caractéristiques de « scalabilité » de la transformée en ondelettes, la notion de « scalabilité » étant définie à

partir de l'expression en anglais « scalability » qui caractérise un encodeur capable d'encoder ou un décodeur capable de décoder un ensemble ordonné de flux numériques de façon à produire ou reconstituer une séquence multi  
5 couches, plusieurs variantes du procédé d'embrouillage et de désembrouillage sont mises en œuvre et sont illustrées avec des exemples de réalisation.

L'invention concerne selon son acception la plus  
10 générale un procédé pour la distribution sécurisée de séquences vidéos selon un format de flux numérique issu d'un encodage basé sur un traitement par ondelettes, constitué de trames (frames) comprenant des blocs contenant des coefficients d'ondelettes décrivant les  
15 éléments visuels, caractérisé en ce que l'on procède, avant la transmission à l'équipement client, à une analyse du flux pour générer un flux principal modifié, par suppression et remplacement de certaines informations codant le flux original et présentant le format du flux  
20 original, et une information complémentaire d'un format quelconque, comportant lesdites informations numériques codant le flux original aptes à permettre la reconstruction desdites trames modifiées, puis à transmettre séparément ledit flux principal modifié et  
25 ladite information complémentaire ainsi générés depuis le serveur vers l'équipement destinataire.

La protection est effectuée par suppression des éléments originaux et leur substitution par des leurres, les éléments originaux extraits étant stockés séparément  
30 dans l'information complémentaire. Le fait d'avoir enlevé et substitué une partie des données originales du flux vidéo d'origine lors de la génération du flux principal modifié, ne permet pas la restitution dudit flux d'origine à partir des seules données dudit flux principal modifié.

Le flux vidéo est entièrement protégé (toutes les sous-bandes) et entièrement transmis via le réseau ou via un support physique à l'utilisateur, indépendamment de ses droits. La reconstitution partielle est effectuée via  
5 l'envoi d'une partie de l'information complémentaire, contenant les éléments originaux et cela directement ou en mode progressif.

Le module d'analyse et d'embrouillage décide comment dégrader visuellement le flux vidéo en fonction de sa  
10 structure et de ses propriétés de scalabilité résultant de la transformée en ondelettes. L'étude se porte sur l'impact de la modification de différentes parties du flux (coefficients, sous-bandes, couches de scalabilité, zones d'intérêt) sur la dégradation visuelle.

15 De préférence, l'embrouillage est effectué en modifiant des coefficients ondelettes appartenant à au moins une sous-bande temporelle résultant de l'analyse temporelle.

Avantageusement, l'embrouillage est effectué en  
20 modifiant des coefficients d'ondelettes appartenant à moins une sous-bande spatiale résultant de l'analyse spatiale d'une sous-bande temporelle.

Avantageusement, l'embrouillage est effectué en modifiant des coefficients d'ondelettes appartenant à  
25 moins une sous-bande temporelle résultant d'une analyse temporelle d'une sous-bande spatiale.

Avantageusement les coefficients ondelettes à modifier sont choisis selon des lois aléatoires et/ou définies a priori.

30 Selon un mode de réalisation particulier, les paramètres pour l'embrouillage dépendent des propriétés de scalabilité temporelle et/ou de scalabilité spatiale et/ou de scalabilité qualitative et/ou de scalabilité en débit et/ou de scalabilité par régions d'intérêt offertes par

les flux numériques générés par les codeurs basés ondelettes.

Avantageusement, l'intensité visuelle de la dégradation des séquences vidéo obtenue est déterminée par la quantité de coefficients d'ondelettes modifiés dans chaque sous-bande spatio-temporelle.

Avantageusement, l'intensité de la dégradation visuelle des séquences vidéos décodées à partir du flux principal modifié est fonction de la position au sein du flux numérique original des données modifiées, lesdites données représentant, selon leurs positions, les valeurs quantifiées selon des précisions différentes des coefficients ondelettes appartenant à une sous-bande spatio-temporelle.

Avantageusement, l'intensité de la dégradation visuelle des séquences vidéo décodées à partir du flux principal modifié est déterminée selon l'appartenance à la couche de qualité des coefficients d'ondelettes modifiés dans chaque sous-bande spatio-temporelle.

Selon un mode de mise en œuvre particulier, la modification des coefficients d'ondelettes est effectuée directement dans le flux binaire.

Selon une variante, la modification des coefficients d'ondelettes est effectuée avec un décodage partiel.

Selon une autre variante, la modification des coefficients ondelettes est effectuée pendant le codage ou en effectuant un décodage puis un ré-encodage complet.

Avantageusement, la taille du flux principal modifié est strictement identique à la taille du flux vidéo numérique original.

Selon une autre variante, la substitution des coefficients d'ondelettes est effectuée avec des valeurs aléatoires ou calculées.

De préférence, la durée de l'embrouillage visuel obtenu au sein d'un groupe de trames est déterminée en fonction de la sous-bande temporelle à laquelle appartiennent les coefficients d'ondelettes modifiés.

5           Avantageusement, l'embrouillage visuel obtenu au sein d'un groupe de trames est limité spatialement dans une région d'intérêt de chaque trame.

De plus, l'information complémentaire est organisée en couches de scalabilité temporelle et/ou spatiale et/ou  
10 qualitative et/ou de débit et/ou sur zone d'intérêt.

Dans une variante, le flux est désembrouillé progressivement avec différents niveaux de qualité et/ou résolution et/ou frame-rate et/ou selon une région d'intérêt via l'envoi de certaines parties de  
15 l'information complémentaire correspondant aux couches de scalabilité respectivement qualitative et/ou spatiale et/ou temporelle et/ou pour une région d'intérêt.

Selon une autre variante, le flux est désembrouillé partiellement selon différents niveaux de qualité et/ou  
20 résolution et/ou frame-rate et/ou selon une région d'intérêt via l'envoi d'une partie de l'information complémentaire correspondant à la (aux) couche(s) de scalabilité respectivement qualitative et/ou spatiale et/ou temporelle et/ou pour cette région d'intérêt.

25           Selon un mode de mise en œuvre particulier, on calcule sur l'équipement destinataire une synthèse d'un flux numérique au format original en fonction dudit flux principal modifié et de ladite information complémentaire.

Selon un mode de mise en œuvre particulier, la  
30 transmission dudit flux principal modifié est réalisée à travers un support matériel distribué physiquement [CD-ROM, DVD, disque dur, carte à mémoire flash].

Avantageusement, le flux principal modifié subit des opérations de transcodage, de réarrangement et/ou

d'extraction de trames ou de groupes de trames pendant sa transmission.

Avantageusement, la transmission de ladite information complémentaire est réalisée à travers un support matériel distribué physiquement [carte à mémoire flash, carte à puce].

De préférence, la modification des coefficients d'ondelettes est parfaitement réversible (procédé sans perte) et le flux numérique reconstitué à partir du flux principal modifié et de l'information complémentaire est strictement identique au flux original.

Avantageusement, la modification des coefficients d'ondelettes est parfaitement réversible (procédé sans perte) et la portion du flux numérique reconstituée à partir du flux principal modifié et de l'information complémentaire est strictement identique à la portion correspondante dans le flux original.

Selon une variante particulière, la reconstitution d'un flux vidéo désemprouillé est contrôlée et/ou limitée en termes de frame-rate et/ou de résolution et/ou de débit et/ou de qualité prédéfinis, en fonction des droits de l'utilisateur.

Selon une autre variante, la reconstitution d'un flux vidéo désemprouillé est limitée en termes de frame-rate et/ou de résolution et/ou de débit et/ou de qualité en fonction de l'appareil de visualisation sur lequel il est visualisé.

Selon une autre variante, la reconstitution du flux vidéo désemprouillé est effectuée de manière progressive par étapes jusqu'à la reconstitution du flux vidéo original.

L'invention concerne également un système pour la fabrication d'un flux vidéo, comportant au moins un serveur multimédia contenant les séquences vidéos

originales et comportant un dispositif d'analyse du flux vidéo, un dispositif de séparation du flux vidéo original en un flux principal modifié par suppression et remplacement de certaines informations codant le signal visuel original et en une information complémentaire en fonction de ladite analyse, et au moins un dispositif sur l'équipement destinataire pour la reconstruction du flux vidéo en fonction dudit flux principal modifié et de ladite information complémentaire.

10

La présente invention sera mieux comprise à la lecture de la description d'un exemple non limitatif de réalisation qui suit, se référant à la figure décrivant l'architecture d'ensemble d'un système pour la mise en œuvre du procédé selon l'invention.

15

La protection des flux visuels décrite est élaborée en se basant sur la structure des flux binaires et leurs caractéristiques dues à l'encodage basé ondelettes. Nous rappelons ci-après cette structure.

20

Un codeur vidéo basé sur un traitement par ondelettes réalise une décomposition temporelle et spatiale d'une séquence vidéo initiale afin d'obtenir un ensemble de coefficients d'ondelettes spatio-temporels. Ces coefficients sont ensuite quantifiés puis codés par un codeur entropique afin de générer un ou plusieurs flux binaires emboîtés qui possèdent des propriétés de scalabilité temporelle et/ou scalabilité spatiale (ou de résolution) et/ou scalabilité qualitative et/ou scalabilité de débit et/ou scalabilité de régions d'intérêt.

25  
30

La propriété de scalabilité temporelle est la possibilité de décoder à partir d'un unique ou plusieurs flux binaires emboîtés, des séquences vidéos dont la

fréquence d'affichage des trames (nombre de trames par seconde) est variable.

Dans la littérature bien connue par l'homme de l'art, sont employées les notions en anglais « frames »  
5 (trames) et « frame rate » (nombre de trames par seconde), notions qui seront utilisées par la suite pour décrire la présente invention.

La propriété de scalabilité spatiale est la possibilité de décoder à partir d'un unique ou à partir de  
10 plusieurs flux binaires emboîtés des séquences vidéos dont la résolution spatiale (taille) des frames est variable.

La propriété de scalabilité qualitative est la possibilité de décoder à partir d'un unique ou plusieurs flux binaires emboîtés des séquences vidéos dont la  
15 qualité visuelle des frames, mesurée selon des critères objectifs et/ou subjectifs, est variable.

La propriété de scalabilité de débit est la possibilité de décoder à partir d'un unique ou plusieurs flux binaires emboîtés des séquences vidéos selon un débit  
20 moyen (nombre moyen de bits d'information par seconde) variable.

La propriété de scalabilité sur région d'intérêt est la possibilité de décoder à partir d'un unique ou plusieurs flux binaires emboîtés une ou plusieurs zones  
25 ciblées au sein de la séquence vidéo.

A l'encodage, une séquence vidéo originale est segmentée en groupes de N frames successives appelé GOF (Group Of Frames), chaque GOF étant ensuite traité de manière indépendante durant l'encodage. On note un GOF de  
30 longueur  $N$   $GOF = (F_0, F_1, \dots, F_{N-1})$ ,  $F_i$  étant les frames pour  $i = 0, 1, 2, \dots, N-1$ . Les coefficients d'ondelettes spatio-temporels sont générés en deux étapes successives et selon une analyse spatiale et temporelle des frames du GOF.



La première étape consiste à effectuer une analyse temporelle des N frames de chaque GOF selon la direction estimée du mouvement (analyse temporelle avec estimation de mouvement), ceci afin de supprimer les redondances temporelles et de concentrer spatialement dans les frames issues de l'analyse temporelle l'énergie et les informations dues au mouvement. Cette analyse temporelle peut être réalisée selon des résolutions spatiales différentes et après décomposition en ondelettes de chaque frame du GOF, l'estimation de mouvement étant réalisée indépendamment au sein de chaque sous-bande spatiale (estimation de mouvement multi-résolution).

La seconde étape consiste à effectuer une analyse spatiale des N frames résultant de l'analyse temporelle à l'aide d'une décomposition en ondelettes afin de supprimer les redondances spatiales et de concentrer l'énergie due aux discontinuités spatiales présentes dans chaque frame.

L'analyse temporelle est réalisée en plusieurs itérations.

A la première itération, p frames successives du GOF original sont analysées au moyen de filtres ondelettes  $f_L$  et  $f_H$  prédéfinis de longueur p et après estimation et compensation du mouvement pour chaque frame et par rapport à une ou plusieurs frames dites frames de référence. Généralement  $p=2$  et deux frames successives ( $F_{2i}, F_{2i+1}$ ),  $i=0, \dots, N/2$ , sont filtrées et engendrent une frame dite de « basse fréquence » ou « moyenne » et notée  $L_i = f_L(F_{2i}, F_{2i+1})$  et une frame dite de « haute fréquence » ou « différence » et notée  $H_i = f_H(F_{2i}, F_{2i+1})$ . A la première itération de l'étape d'analyse temporelle sont ainsi générés un sous-ensemble t- $L_1$  de frames de type L de longueur  $N/2$  et un sous-ensemble t- $H_1$  de frames de type H de longueur  $N/2$ , tels que

$$t-L_1 = (L_0, L_1, \dots, L_{N/2}),$$

$$t-H_1 = (H_0, H_1, \dots, H_{N/2}).$$

A chaque itération  $k > 1$  suivante, l'estimation/compensation de mouvement et le filtrage temporel sont itérés sur le sous-ensemble de frames  $t-L_{k-1}$  de type L obtenues à l'itération  $k-1$  et sont générés deux nouveaux sous-ensembles de frames  $t-L_k$  et  $t-H_k$  dont la longueur (nombre de frames) est diminuée d'un facteur 2 par rapport à  $t-L_{k-1}$ . Dans certains cas, l'itération porte aussi sur le sous-ensemble de frames  $t-H_{k-1}$ .

Le nombre total d'itérations durant l'analyse temporelle est noté  $n_T$ . Il est compris entre 1 et  $N/2$ . A l'issue de l'analyse temporelle,  $n_T+1$  sous-bandes temporelles sont générées.

15

Par exemple, avec  $N=16$  et  $n_T=4$  :

$$L_i^k = f_L(L_{2i}^{k-1}, L_{2i+1}^{k-1}),$$

$$H_i^k = f_H(L_{2i}^{k-1}, L_{2i+1}^{k-1})$$

GOF = (F <sub>0</sub> , F <sub>1</sub> , ..., F <sub>15</sub> )			
t-L <sub>1</sub> = (L <sub>0</sub> , ..., L <sub>7</sub> )			t-H <sub>1</sub> = (H <sub>0</sub> , ..., H <sub>7</sub> )
t-L <sub>2</sub> = (LL <sub>0</sub> , ..., LL <sub>3</sub> )		t-H <sub>2</sub> = (LH <sub>0</sub> , ..., LH <sub>3</sub> )	
t-L <sub>3</sub> = (LLL <sub>0</sub> , LLL <sub>1</sub> )	t-H <sub>3</sub> = (LLH <sub>0</sub> , LLH <sub>1</sub> )		
t-L <sub>4</sub> = (LLLL <sub>0</sub> )	t-H <sub>4</sub> = (LLLH <sub>0</sub> )		

20

L'analyse temporelle d'un GOF de longueur  $N=16$  avec  $n_T=4$  génère ainsi 16 nouvelles frames réparties dans  $n_T+1=4+1=5$  sous-bandes temporelles :

- Sous-bande t-L<sub>4</sub> : 1 frame de type L : LLLL<sub>0</sub>,
- Sous-bande t-H<sub>4</sub> : 1 frame de type H : LLLH<sub>0</sub>,
- Sous-bande t-H<sub>3</sub> : 2 frames de type H : LLH<sub>0</sub>, LLH<sub>1</sub>,
- Sous-bande t-H<sub>2</sub> : 4 frames de type H : LH<sub>0</sub>, ..., LH<sub>3</sub>

25

- Sous-bande  $t-H_1$  : 8 frames de type H :  $H_0, \dots, H_7$ .

L'analyse spatiale est réalisée ensuite sur chacune des frames appartenant à chaque sous-bande temporelle  $t-L_i$  et  $t-H_i$  : chaque frame est décomposée à l'aide d'une transformée en ondelettes discrète à D niveaux, générant ainsi  $3D+1$  sous-bandes spatiales de coefficients d'ondelettes pour chaque frame. Ces sous-bandes spatiales sont notées  $s-LL_0$ ,  $s-HL_1$ ,  $s-LH_1$ ,  $s-HH_1$ ,  $s-HL_2$ ,  $s-LH_2$ ,  $s-HH_2$ , ...,  $s-HL_D$ ,  $s-LH_D$ ,  $s-HH_D$ .

10 A l'issue des analyses temporelles et spatiales,  $(n_T+1) \times (3D+1)$  sous-bandes spatio-temporelles de coefficients d'ondelettes sont disponibles :

- $t-L_{n_T}(s-LL_0)$ ,  $t-L_{n_T}(s-HL_1)$ ,  $t-L_{n_T}(s-LH_1)$ ,  $t-L_{n_T}(s-HH_1)$ , ...,  $t-L_{n_T}(s-HL_D)$ ,  $t-L_{n_T}(s-LH_D)$ ,  $t-L_{n_T}(s-HH_D)$ ,

- 15 ▪  $t-H_{n_T}(s-LL_0)$ ,  $t-H_{n_T}(s-HL_1)$ ,  $t-H_{n_T}(s-LH_1)$ ,  $t-H_{n_T}(s-HH_1)$ , ...,  $t-H_{n_T}(s-HL_D)$ ,  $t-H_{n_T}(s-LH_D)$ ,  $t-H_{n_T}(s-HH_D)$ ,

- ...

$t-H_1(s-LL_0)$ ,  $t-H_1(s-HL_1)$ ,  $t-H_1(s-LH_1)$ ,  $t-H_1(s-HH_1)$ , ...,  $t-H_1(s-HL_D)$ ,  $t-H_1(s-LH_D)$ ,  $t-H_1(s-HH_D)$ .

20 Les coefficients ondelettes de chaque sous-bande spatio-temporelle sont ensuite compressés progressivement par plan de bits à l'aide d'un codeur entropique dont le but est de supprimer les redondances statistiques existant au sein d'un ensemble fixé de coefficients ondelettes. Le

25 codeur entropique génère, pour chaque ensemble de coefficients ondelettes codés indépendamment, un flux binaire qui peut être découpé en plusieurs sous-flux répartis selon différentes couches de qualité.

30 Après une analyse de la structure en sous-bandes décrite précédemment, le module d'analyse et d'embrouillage conforme à l'invention effectue des modifications (par permutation et/ou substitution et/ou seuillage) d'un sous-ensemble des coefficients

d'ondelettes appartenant à une ou plusieurs sous-bandes spatio-temporelles. Ces modifications introduisent une dégradation perceptible visuellement (embrouillage) de la séquence vidéo décodée à partir de ces coefficients modifiés. En fonction du nombre de coefficients modifiés, de leurs localisations dans une sous-bande spatiale, de leurs appartenances aux sous-bandes spatio-temporelles, de leurs appartenances à une ou plusieurs couches de qualité, de leur position au sein de l'ensemble des coefficients appartenant à une seule sous-bande spatio-temporelle et du type de modification, un contrôle de l'étendue spatiale et/ou temporelle et/ou selon les couches de qualité de l'embrouillage, ainsi qu'un contrôle de l'intensité de la dégradation due à l'embrouillage sont possibles.

En modifiant des coefficients d'ondelettes dans des sous-bandes spatiales d'une sous-bande temporelle  $t$ - $X$  de longueur  $l_x$  (i.e. contenant  $l_x$  frames), les  $N/l_x$  frames consécutives du GOF sont embrouillées.

Le choix du type de sous-bande spatiale à laquelle appartient les coefficients d'ondelettes (s-HL ou s-LH ou s-HH) permet de contrôler l'aspect visuel de l'embrouillage : pour la sous-bande s-HL des artefacts de direction verticale apparaissent sur les frames (dégradation des discontinuités spatiales verticales), pour la sous-bande s-LH des artefacts horizontaux apparaissent (dégradation des discontinuités spatiales horizontales) et pour la sous-bande s-HH des artefacts de type « damier » apparaissent (dégradations conjointes des discontinuités spatiales horizontales et verticales).

Le choix du niveau de résolution  $r$  auquel appartient la sous-bande spatiale (s-LL <sub>$r$</sub>  ou s-HL <sub>$r$</sub>  ou s-LH <sub>$r$</sub>  ou s-HH <sub>$r$</sub> ) permet de contrôler l'étendue spatiale de l'embrouillage engendré par la modification des coefficients

d'ondelettes : plus  $r$  est proche de 0, plus l'étendue spatiale sera importante.

Une modification des coefficients d'ondelettes appartenant à une sous-bande de résolution  $r \geq 0$  génère un embrouillage qui sera visible sur toutes les frames  
5 décodées de résolutions spatiales supérieures  $r+1$ ,  $r+2$ , ...,  $R$ .

Une modification des coefficients d'ondelettes appartenant à une couche de qualité  $q$  génère un  
10 embrouillage qui sera visible sur toutes les frames décodées en considérant au moins les  $q$  premières couches de qualité.

La modification des coefficients d'ondelettes spatio-temporels est réalisée après un décodage partiel du  
15 flux binaire généré selon un standard ou une norme ou un algorithme ou un format d'encodage. Une fois la modification effectuée, un ré-encodage des coefficients est réalisé afin de générer un flux binaire de taille identique et qui respecte la conformité vis-à-vis du  
20 standard ou de la norme ou de l'algorithme ou du format d'encodage ayant généré le flux binaire original.

Avantageusement, des sous-ensembles de bits à l'intérieur même du flux binaire original représentant les coefficients d'ondelettes spatio-temporels codés sont  
25 modifiés sans décodage et sans perturber la conformité du flux vis-à-vis du standard ou de la norme ou de l'algorithme ou du format d'encodage ayant généré le flux binaire original.

Le choix des coefficients d'ondelettes spatio-temporels à modifier au sein d'une sous-bande spatio-temporelle est conduit de manière aléatoire et/ou selon des règles définies a priori.  
30

Avantageusement, le flux principal modifié a une taille identique à celle du flux vidéo original.

L'embrouillage ainsi généré présente des propriétés de scalabilité temporelle, spatiale, qualitative, de débit et sur zone d'intérêt.

Avantageusement, l'information complémentaire  
5 relative audit embrouillage ainsi généré est organisée en couches de scalabilité temporelle, spatiale, qualitative, de débit et sur zone d'intérêt.

En fonction du nombre de GOF et/ou du nombre de frames embrouillées au sein d'un GOF, l'embrouillage  
10 présente une scalabilité temporelle comprise entre : « toutes les frames de tous les GOF (embrouillage maximal) » et « aucune frame d'aucun GOF » (séquence non embrouillée).

En fonction des résolutions des sous-bandes spatiales auxquelles appartiennent les coefficients d'ondelettes modifiés, l'embrouillage présente une  
15 scalabilité spatiale comprise entre : « toutes les résolutions sont embrouillées » (i.e. de la résolution  $r=0$  à la résolution  $r=R$ ) et « aucune des résolutions n'est embrouillée ».  
20

En fonction du nombre de coefficients d'ondelettes modifiés et des résolutions des sous-bandes spatiales auxquelles ils appartiennent, l'embrouillage présente une scalabilité qualitative s'échelonnant de : « la totalité  
25 de chaque frame est embrouillée », « certaines régions spatiales de chaque frame sont embrouillées » (régions d'intérêts) et « aucun embrouillage n'est appliqué sur les frames ».

Réciproquement, le désembrouillage présentera lui  
30 aussi les différentes scalabilités énoncées (temporelle, spatiale, qualitative, de débit et de zone d'intérêt).

Avantageusement, ledit désembrouillage permettra d'adresser les différentes scalabilités énoncées (temporelle, spatiale, qualitative, de débit et de zone

d'intérêt), grâce à l'envoi de certaines parties de l'information complémentaire correspondant à différentes couches de scalabilité (temporelle, spatiale, qualitative, de débit et de zone d'intérêt), donnant ainsi accès à des  
5 niveaux différents de qualité/résolution/frame-rate pour la séquence vidéo décodée à partir du flux désembrouillé partiellement.

Avantageusement, les différents niveaux de qualité/résolution/frame rate de la séquence vidéo sont  
10 obtenus à partir du flux désembrouillé partiellement via l'envoi d'une partie de l'information complémentaire par couche de scalabilité (temporelle, spatiale, qualitative, de débit et de zone d'intérêt).

On comprendra mieux le principe d'embrouillage et de  
15 désembrouillage se basant sur ces différentes scalabilités à l'aide de l'exemple non limitatif de réalisation préféré qui suit.

Sur le dessin en annexe, la figure représente un mode de réalisation préféré particulier du système client-  
20 serveur conforme à l'invention.

Le flux d'origine (1) est directement sous forme numérique ou sous forme analogique. Dans ce dernier cas, le flux analogique est converti par un codeur basé ondelettes non représenté en un format numérique (2). Le  
25 flux vidéo que l'on souhaite sécuriser (2) est passé à un module d'analyse et d'embrouillage (3) qui va générer un flux principal modifié (5) au format identique au flux d'entrée (2) en dehors de ce que certains des coefficients ont été remplacés par des valeurs différentes de celles  
30 d'origine, et est stocké dans le serveur (6). L'information complémentaire (4), de format quelconque, est également placée dans le serveur (6) et contient des informations relatives aux éléments des images qui ont été

modifiés, remplacés, substitués ou déplacés, et leur valeur ou emplacement dans le flux original.

Le flux (5) au format identique au flux original est ensuite transmis, via un réseau haut débit (9) de type  
5 hertzien, câble, satellite, etc., au terminal du spectateur (8), et plus précisément sur son disque dur (10). Lorsque le spectateur (8) fait la demande de visionnage du film présent sur son disque dur (10), deux éventualités sont possibles : soit le spectateur (8) ne  
10 possède pas tous les droits nécessaires pour voir le film, dans ce cas, le flux vidéo (5) généré par le module d'embrouillage (3) présent sur le disque dur (10) est passé au système de synthèse (13), via une mémoire tampon de lecture (11), qui ne le modifie pas et le transmet à  
15 l'identique à un lecteur afficheur capable de le décoder (14) et son contenu, dégradé visuellement par le module d'embrouillage (3), est affiché sur l'écran de visualisation (15). Avantageusement, le flux vidéo (5) généré par le module d'embrouillage (3) est passé  
20 directement via un réseau (9) à la mémoire tampon de lecture (11) puis au système de synthèse (13).

Avantageusement, le flux vidéo (5) subit dans le réseau (9) une série d'opérations de transcodage et de réarrangement de ses trames ou groupes de trames.

25 Soit le serveur décide que le spectateur (8) possède les droits pour visualiser correctement le film. Dans ce cas, le module de synthèse (13) fait une demande de visionnage au serveur (6) contenant l'information complémentaire nécessaire (4) à la reconstitution de la  
30 vidéo originale (2). Le serveur (6) envoie alors via le réseau de télécommunication (7) de type ligne téléphonique analogique ou numérique, DSL (Digital Subscriber Line) ou BLR (Boucle Locale Radio), via des réseaux DAB (Digital Audio Broadcasting) ou via des réseaux de



télécommunications mobiles numériques (GSM, GPRS, UMTS), l'information complémentaire (4), permettant la reconstitution de la vidéo originale, de façon à ce que le spectateur (8) puisse la stocker dans une mémoire tampon (12). Le module de synthèse (13) procède alors à la restauration du flux vidéo embrouillé qu'il lit dans sa mémoire tampon de lecture (11), des champs modifiés dont il connaît les positions ainsi que les valeurs d'origine sont restituées, grâce au contenu de l'information complémentaire lue dans la mémoire tampon (12) de désembrouillage. La quantité d'informations contenue dans l'information complémentaire (4) qui est envoyée au module de désembrouillage est spécifique, adaptative et progressive pour chaque spectateur et dépend de ses droits, par exemple utilisation unique ou multiple, droit de faire une ou plusieurs copies privées, retard ou anticipation de paiement. Le niveau (qualité, quantité, type) de l'information complémentaire est déterminé également en fonction de la qualité visuelle requise par l'utilisateur. Le codage vidéo basé ondelettes se caractérisant par les scalabilités précédemment décrites permet la restitution du flux vidéo avec des niveaux de qualité, de résolution, de fréquence de trames différents.

Avantageusement, le flux principal modifié (5) est passé directement via un réseau (9) à la mémoire tampon de lecture (11) puis au module de synthèse (13).

Avantageusement, le flux principal modifié (5) est inscrit (enregistré) sur un support physique comme un disque de type CD-ROM, DVD, disque dur, carte à mémoire flash, etc. (9bis). Le flux principal modifié (5) sera ensuite lu depuis le support physique (9bis) par le lecteur de disque (10bis) du boîtier (8) pour être transmis à la mémoire tampon de lecture (11) puis au module de synthèse (13).

Avantageusement, l'information complémentaire (4) est enregistrée sur un support physique (7bis) de format carte de crédit, constitué par une carte à puce ou une carte à mémoire flash. Cette carte (7bis) sera lue par le module (12) du dispositif (8) qui comprend un lecteur de carte (7ter).

Avantageusement, la carte (7bis) contient les applications et les algorithmes qui seront exécutés par le système de synthèse (13).

Avantageusement, le dispositif (8) est un système autonome, portable et mobile.

Nous allons détailler le fonctionnement du module d'analyse et d'embrouillage (3) illustrant le choix de l'embrouillage effectué. La séquence vidéo originale est segmentée en GOF de  $N=16$  frames. L'analyse temporelle avec  $n_T=4$  itérations génère  $n_T+1=5$  sous-bandes temporelles possédant respectivement :

- Sous-bande  $t-L_4$  : 1 frame de type L :  $LLLL_0$
- Sous-bande  $t-H_4$  : 1 frame de type H :  $LLLH_0$
- Sous-bande  $t-H_3$  : 2 frames de type H :  $LLH_0, LLH_1$
- Sous-bande  $t-H_2$  : 4 frames de type H :  $LH_0, LH_1, LH_2, LH_3$
- Sous-bande  $t-H_1$  : 8 frames de type H :  $H_0, H_1, H_2, H_3, H_4, H_5, H_6, H_7$ .

La décomposition en cinq sous-bandes temporelles offre la possibilité de reconstituer la séquence vidéo initiale selon cinq différents « frame rate ». Chaque frame de résolution  $R$  dans chaque sous-bande temporelle  $t-X$  est ensuite décomposée spatialement par une transformée en ondelettes à  $D=4$  niveaux, ce qui donne la possibilité de reconstituer l'image avec cinq résolutions différentes, générant ainsi pour chacune  $3 \times D + 1 = 13$  sous-bandes spatiales :  $LL_0, LH_1, HL_1, HH_1, LH_2, HL_2, HH_2, LH_3, HL_3, HH_3, LH_4, HL_4, HH_4$ .

Suite à un tel encodage, la séquence vidéo peut donc être décodée selon des frame-rates allant de  $1/16 \times fr_0$  à  $fr_0$ ,  $fr_0$  étant le frame-rate de la séquence vidéo originale ainsi que selon  $D+1=5$  résolutions.

5 L'embrouillage de la séquence vidéo est réalisée pour chaque GOF de la manière suivante :

- Dans la sous-bande temporelle  $t-L_4$ , les coefficients d'ondelettes des sous-bandes spatiales  $s-HH_2$  et  $s-HH_3$  résultant de la décomposition spatiale en  
10 ondelettes de la frame  $LLLL_0$  sont extraits et remplacés par des valeurs aléatoires ou calculées.

- Dans la sous-bande temporelle  $t-H_3$ , les coefficients d'ondelettes de la sous-bande spatiale  $s-HH_3$  résultant de la décomposition spatiale en ondelettes de la  
15 frame  $LLH_0$  sont extraits et remplacés par des valeurs aléatoires ou calculées.

- Dans la sous-bande temporelle  $t-H_1$ , les coefficients d'ondelettes de la sous-bande spatiale  $s-HH_3$  résultant de la décomposition en ondelettes de la frame  
20  $LH_0$  sont extraits et remplacés par des valeurs aléatoires ou calculées.

- Dans la sous-bande temporelle  $t-H_0$ , les coefficients d'ondelettes de la sous-bande spatiale  $s-HH_3$  résultant de la décomposition en ondelettes de la frame  $H_0$   
25 sont extraits et remplacés par des valeurs aléatoires ou calculées.

La séquence vidéo décodée à partir du flux principal modifié est donc embrouillée totalement, sauf si elle est décodée à des frame-rates égaux à  $1/16 \times fr_0$  et  $1/8 \times fr_0$   
30 (décodage uniquement à partir respectivement des sous-bandes temporelles  $t-L_4$  et  $(t-L_4, t-H_4)$ ) et à des résolutions spatiales  $r=R/16$  ou  $r=R/8$  (décodage uniquement à partir respectivement des sous-bandes spatiales  $s-LL_0$ ,  $s-LH_1$ ,  $s-HL_1$  et  $s-HH_1$  qui n'ont pas été modifiées). On

effectue ainsi des modifications dans toutes les bandes temporelles, mais pas à toutes les résolutions. Le fait de laisser des résolutions non modifiées permet de pouvoir reconstituer le flux vidéo à partir du flux embrouillé, 5 mais à une qualité nettement inférieure que celle du flux vidéo original.

Le flux ainsi embrouillé est transmis au client (8) à sa demande, le désembrouillage est ensuite réalisé par exemple en cinq étapes, correspondant à différents 10 niveaux de qualité obtenus après chaque étape de désembrouillage. De cette manière, on effectue un désembrouillage d'une ou plusieurs couches de scalabilité, la qualité du film visualisé étant contrôlée par le serveur en fonction des droits de l'utilisateur et de la 15 qualité requise par lui.

Avantageusement, ce désembrouillage se traduit par une atténuation progressive de la dégradation dans le temps jusqu'à reconstituer le contenu original de haute qualité visuelle. Par exemple, la première étape de 20 désembrouillage consiste à restituer les coefficients d'ondelettes de la sous-bande spatiale  $s-HH_2$  pour la sous-bande temporelle  $t-L_4$ . L'embrouillage de la séquence vidéo décodée pour une résolution maximale  $R$  et à un frame-rate maximal de  $fr_0$  est alors moins étendu spatialement et plus 25 concentré autour des discontinuités spatiales de chaque frame. La séquence vidéo décodée pour les frame-rate  $1/16 \times fr_0$  ou  $1/8 \times fr_0$  et pour les résolutions  $r=R/16$ ,  $R/8$ ,  $R/4$  n'est en revanche pas du tout embrouillée. La qualité visuelle du film partiellement désembrouillé est minimale, 30 le film est inexploitable visuellement à sa pleine résolution et à son frame-rate original. Cette étape sert d'identification du serveur d'information complémentaire lors de l'établissement de la connexion.

La deuxième étape de désembrouillage consiste à restituer les coefficients d'ondelettes de la sous-bande spatiale  $s-HH_3$  pour la sous-bande temporelle  $t-L_4$ . L'embrouillage de la séquence vidéo décodée pour une

5 résolution  $R$  et à un frame-rate de  $fr_0$  est alors beaucoup moins étendu spatialement et plus concentré autour des discontinuités spatiales de chaque frame. De plus, la séquence vidéo n'est maintenant embrouillée que sur une durée équivalente à la moitié d'un GOF (8 frames sur 16).

10 La séquence vidéo décodée pour les frame-rate  $1/16 \times fr_0$  ou  $1/8 \times fr_0$  et pour les résolutions  $r=R/16$ ,  $R/8$ ,  $R/4$  n'est pas embrouillée. Cette étape sert à ce que l'utilisateur (8) aperçoive le film vidéo partiellement, afin de décider s'il veut obtenir des droits pour regarder le film. Après

15 confirmation du client, et en fonction du paiement effectué, sont exécutées les étapes trois, quatre ou cinq. La troisième étape de désembrouillage consiste à restituer les coefficients d'ondelettes de la sous-bande spatiale  $s-HH_3$  pour la sous-bande temporelle  $t-H_3$ . La séquence vidéo

20 n'est maintenant embrouillée que sur une durée équivalente au quart d'un GOF (4 frames sur 16). La séquence vidéo décodée pour les frame-rate  $1/16 \times fr_0$ ,  $1/8 \times fr_0$ ,  $1/4 \times fr_0$  et pour les résolutions  $r=R/16$ ,  $R/8$ ,  $R/4$  n'est pas embrouillée. Le film peut être regardé visuellement, mais

25 il est de basse qualité.

La quatrième étape de désembrouillage consiste à restituer les coefficients d'ondelettes de la sous-bande spatiale  $s-HH_3$  pour la sous-bande temporelle  $t-H_2$ . La séquence vidéo n'est maintenant embrouillée que sur une

30 durée équivalente au huitième d'un GOF (2 frames sur 16). La séquence vidéo décodée pour les frame-rate  $1/16 \times fr_0$ ,  $1/8 \times fr_0$ ,  $1/4 \times fr_0$ ,  $1/2 \times fr_0$  et pour les résolutions  $r=R/16$ ,  $R/8$ ,  $R/4$  n'est pas embrouillée. La qualité visuelle du film restitué est moyenne. Pour ces 3<sup>ième</sup> et 4<sup>ième</sup> étapes de

désembrouillage, le film vidéo pleine résolution et pour le frame rate original reste partiellement embrouillé, mais à partir de ces étapes, on peut extraire des flux vidéo avec des résolutions et frame-rate inférieurs à ceux du film original. Cela donne la possibilité de fournir des versions du même flux vidéo de moindre résolution, donc à plus faible prix et de mieux contrôler l'accès.

La cinquième étape de désembrouillage consiste à restituer les coefficients d'ondelettes de la sous-bande spatiale  $s-HH_1$ , pour la sous-bande temporelle  $t-H_1$ . La séquence vidéo est alors totalement désembrouillée, quel que soit le frame-rate de décodage et la résolution. Le flux vidéo reconstitué est strictement identique au flux vidéo original.

## REVENDICATIONS

1. Procédé pour la distribution sécurisée de  
5 séquences vidéos selon un format de flux numérique issu  
d'un encodage basé sur un traitement par ondelettes,  
constitué de trames (frames) comprenant des blocs  
contenant des coefficients d'ondelettes décrivant les  
10 éléments visuels, caractérisé en ce que l'on procède,  
avant la transmission à l'équipement client, à une analyse  
du flux pour générer un flux principal modifié par  
suppression et remplacement de certaines informations  
codant le flux original et présentant le format du flux  
original, et une information complémentaire d'un format  
15 quelconque, comportant lesdites informations numériques  
codant le flux original aptes à permettre la  
reconstruction desdites trames modifiées, puis à  
transmettre séparément ledit flux principal modifié et  
ladite information complémentaire ainsi générés depuis le  
20 serveur vers l'équipement destinataire.

2. Procédé pour la distribution sécurisée de  
séquences vidéos selon la revendication 1, caractérisé en  
ce que l'embrouillage est effectué en modifiant des  
25 coefficients d'ondelettes appartenant à au moins une sous-  
bande temporelle résultant de l'analyse temporelle.

3. Procédé pour la distribution sécurisée de  
séquences vidéos selon l'une des revendications  
30 précédentes, caractérisé en ce que l'embrouillage est  
effectué en modifiant des coefficients d'ondelettes  
appartenant à moins une sous-bande spatiale résultant de  
l'analyse spatiale d'une sous-bande temporelle.

4. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que l'embrouillage est effectué en modifiant des coefficients d'ondelettes appartenant à moins une sous-bande temporelle résultant de l'analyse temporelle d'une sous-bande spatiale.

5. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que les coefficients ondelettes à modifier sont choisis selon des lois aléatoires et/ou définies a priori.

6. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que les paramètres pour l'embrouillage dépendent des propriétés de scalabilité temporelle et/ou de scalabilité spatiale et/ou de scalabilité qualitative et/ou de scalabilité temporelle, de scalabilité en débit et/ou de scalabilité par régions d'intérêt offertes par les flux numériques générés par les codeurs basés ondelettes.

7. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que l'intensité visuelle de la dégradation des séquences vidéo obtenue est déterminée par la quantité de coefficients d'ondelettes modifiés dans chaque sous-bande spatio-temporelle.

30

8. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que l'intensité de la dégradation visuelle des séquences vidéos décodées à



partir du flux principal modifié est fonction de la position au sein du flux numérique original des données modifiées, lesdites données représentant, selon leurs positions, les valeurs quantifiées selon des précisions  
5 différentes des coefficients ondelettes appartenant à une sous-bande spatio-temporelle.

9. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications  
10 précédentes, caractérisé en ce que l'intensité de la dégradation visuelle des séquences vidéo décodées à partir du flux principal modifié est déterminée selon l'appartenance à une couche de qualité des coefficients d'ondelettes modifiés dans chaque sous-bande spatio-  
15 temporelle.

10. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que la modification des  
20 coefficients d'ondelettes est effectuée directement dans le flux binaire.

11. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications  
25 précédentes, caractérisé en ce que la modification des coefficients d'ondelettes est effectuée avec un décodage partiel.

12. Procédé pour la distribution sécurisée de  
30 séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que la modification des coefficients ondelettes est réalisée pendant le codage ou en effectuant un décodage puis un ré-encodage complet.

13. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que la taille du flux principal modifié est strictement identique à la taille du flux vidéo numérique original.

14. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que la substitution des coefficients d'ondelettes est effectuée avec des valeurs aléatoires ou calculées.

15. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que la durée de l'embrouillage visuel obtenu au sein d'un groupe de trames est déterminée en fonction de la sous-bande temporelle à laquelle appartiennent les coefficients d'ondelettes modifiés.

20

16. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que l'embrouillage visuel obtenu au sein d'un groupe de trames est limité spatialement dans une région d'intérêt de chaque trame.

17. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que l'information complémentaire est organisée en couches de scalabilité temporelle et/ou spatiale et/ou qualitative et/ou de débit et/ou sur zone d'intérêt.

30

18. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que le flux est désembrouillé progressivement avec différents niveaux de  
5 qualité et/ou résolution et/ou frame-rate et/ou selon une région d'intérêt via l'envoi de certaines parties de l'information complémentaire correspondant aux couches de scalabilité respectivement qualitative et/ou spatiale et/ou temporelle et/ou pour une région d'intérêt.

10

19. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que le flux est désembrouillé partiellement selon différents niveaux de  
15 qualité et/ou résolution et/ou frame-rate et/ou selon une région d'intérêt via l'envoi d'une partie de l'information complémentaire correspondant à la (aux) couche(s) de scalabilité respectivement qualitative et/ou spatiale et/ou temporelle et/ou pour cette région d'intérêt.

20

20. Procédé pour la distribution sécurisée de séquences audiovisuelles selon l'une des revendications précédentes, caractérisé en ce que l'on calcule sur l'équipement destinataire une synthèse d'un flux numérique  
25 au format original en fonction dudit flux principal modifié et de ladite information complémentaire.

21. Procédé pour la distribution sécurisée de séquences audiovisuelles selon l'une des revendications  
30 précédentes, caractérisé en ce que la transmission dudit flux principal modifié est réalisée à travers un support matériel distribué physiquement [CD-ROM, DVD, disque dur, carte à mémoire flash].

22. Procédé pour la distribution sécurisée de séquences audiovisuelles selon l'une des revendications précédentes, caractérisé en ce que le flux principal modifié subit des opérations de transcodage, de réarrangement et/ou d'extraction de trames ou de groupes de trames pendant sa transmission.

23. Procédé pour la distribution sécurisée de séquences audiovisuelles selon l'une des revendications précédentes, caractérisé en ce que la transmission de ladite information complémentaire est réalisée à travers un support matériel distribué physiquement [carte à mémoire flash, carte à puce].

24. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que la modification des coefficients d'ondelettes est parfaitement réversible (procédé sans perte) et que le flux numérique reconstitué à partir du flux principal modifié et de l'information complémentaire est strictement identique au flux original.

25. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que la modification des coefficients d'ondelettes est parfaitement réversible (procédé sans perte) et que la portion du flux numérique reconstituée à partir du flux principal modifié et de l'information complémentaire est strictement identique à la portion correspondante dans le flux original.

26. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications 24 ou 25, caractérisé en ce que la reconstitution d'un flux vidéo

désembrouillé est contrôlée et/ou limitée en termes de frame-rate et/ou de résolution, et/ou de débit et/ou de qualité prédéfinis en fonction des droits de l'utilisateur.

5

27. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications 24 ou 25, caractérisé en ce que la reconstitution d'un flux vidéo désembrouillé est contrôlée et/ou limitée en terme de  
10 frame-rate et/ou de résolution, et/ou de débit et/ou de qualité prédéfinis en fonction de l'appareil de visualisation sur lequel il est visualisé.

28. Procédé pour la distribution sécurisée de  
15 séquences vidéos selon l'une des revendications 24, 25, 26 ou 27, caractérisé en ce que la reconstitution du flux vidéo désembrouillé est effectuée de manière progressive par étapes jusqu'à la reconstitution du flux vidéo original.

20

29. Système pour la fabrication d'un flux vidéo pour la mise en œuvre du procédé selon l'une des revendications précédentes, comportant au moins un serveur multimédia contenant les séquences vidéos originales et caractérisé  
25 en ce qu'il comporte un dispositif d'analyse du flux vidéo, un dispositif de séparation du flux vidéo original en un flux principal modifié par suppression et remplacement de certaines informations codant le flux visuel original et en une information complémentaire en  
30 fonction de ladite analyse, et au moins un dispositif sur l'équipement destinataire pour la reconstruction du flux vidéo en fonction dudit flux principal modifié et de ladite information complémentaire.

1/1

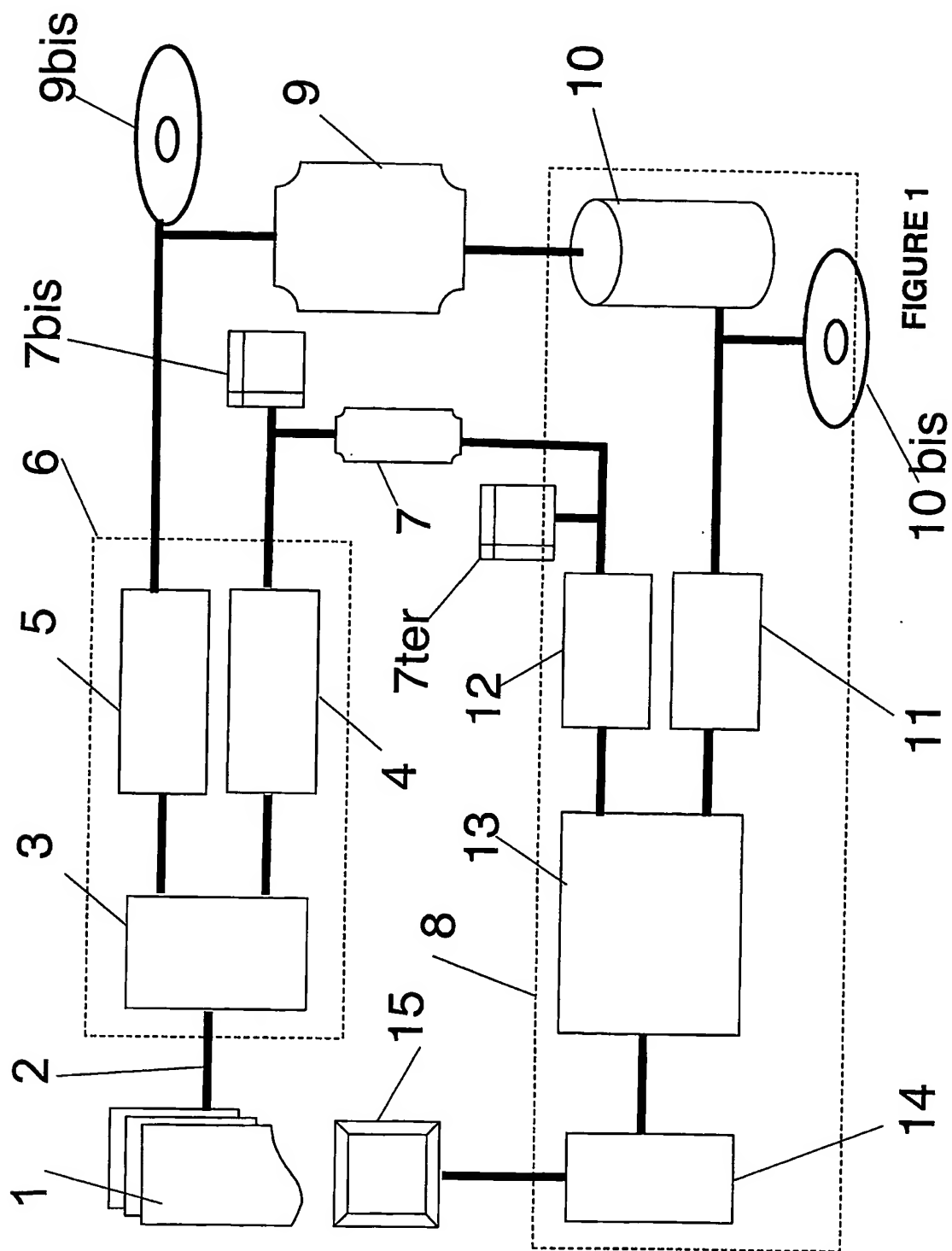


FIGURE 1